STORAGE DEVICE SYSTEM HAVING BI-DIRECTIONAL COPY CONTROL FUNCTION

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a storage device system, and more particularly, relates to a technology for carrying out copying to another storage device system from a certain storage device system in a bi-directional manner.

2. Description of the Related Art

In an information system having a host computer as a superior device and a plurality of storage device systems (storage systems), as a technology for carrying out data copying between storage device systems, there is a remote copy. The remote copy is a copy of data between a plurality of storage device systems which exist at physically remote locations, without intervention of a superior device, in an information processing system, i.e., a technology for carrying out double writing. The storage device system indicates a system which is configured by a plurality of storage devices and a controller for controlling those storage devices. In the information processing system for carrying out the remote copy, storage device systems which are disposed respectively at physically remote locations are mutually connected by a dedicated line or a public line. The dedicated line or the public line which is

used for this connection is called as a remote copy line.

Among logical storage regions (hereinafter, referred to as "logical volume") that a certain storage device system has, a logical volume of the same capacity as capacity of a logical volume which becomes an object of a remote copy (hereinafter, referred to as "copy-source logical volume") is secured in another storage device system for copying the same. This secured logical volume (hereinafter, referred to as "copy-destination logical volume") is formed so as to have a one-to-one correspondence relationship with the copy-source logical volume (hereinafter, referred to as "copy pair"). Data of the copy-source logical volume is copied to the copy-destination logical volume through the dedicated line or the public line. In case that there was updating of data which is included in the copy-source logical volume, updated data is transferred through the dedicated line etc. to a storage device system having the copy-source logical volume, and the updated data was also written into the copy-destination logical volume which corresponds to the copy-source logical volume. Using the remote copy technology, in the information processing system having a plurality of storage device systems, it is possible to hold logical volumes of the same contents in the plurality of storage device systems.

By forming the copy pair of a copy source and a copy destination, a copy direction from the copy source to the copy

destination is determined to become one direction. As to the copy-source logical volume, it is possible to carry out a write-in process from a superior device. Adversely, as to the copy-destination logical volume, it is impossible to carry out a write-in process to a superior device.

Supposing that the copy direction is not fixed to one direction and enabled is write-in to a storage device system which configures the copy pair, it is impossible to hold identical contents of data in respective storage device systems. This is because of transfer time which is required for transfer of copy data when a superior device carries out write-in to a storage device, and thereafter, a content to be written is written into a storage device as the copy destination.

An environment in which the copy pair is formed between a storage device system A and a storage device system B will be concretely described as an example. Here, the storage device system A and the storage device system B are sufficiently away from each other geographically, and it is to take, for example, 1 second and more until data is written into the storage device system A from a superior device, and then, double data is transferred and written into the storage device system B from the storage device system A.

Here, supposing that, at almost the same time, different contents (content A, content B, respectively) were written into the same regions of the storage device system A and the storage

device system B from a superior device, the content A and the content B are written into the storage device system A and the storage device system B, respectively. After respective write-ins were completed, at almost the same timing, from the storage device system A to the storage device system B, also from the storage device system B to the storage device system A, the content A and the content B are transferred. Under the suchlike situation, after the storage device system A and the storage device system B received data transferred, the content A and the content B are to have been overwritten in the storage device system A and the storage device system B, respectively. In the storage device system A, realized is such a situation that the content B was overwritten on the content A, and in the storage device system B, realized is such a situation that the content A was overwritten on the content B. In the suchlike case, contents which were written into the storage device system A and the storage device system B become different ones, which results in that gemination of the volume has not been carried out.

In order to avoid the suchlike situation and to realize complete gemination of the volume, a copy direction is decided to be one direction like the copy source and the copy destination. A technology regarding the remote copy is disclosed in U.S patent No.5,742,792 (Patent Document 1).

Up to now, a storage device which is shared by a plurality

of superior devices, on the basis of a shared exclusive control request from an arbitrary superior device, realizes a shared exclusive control to an access request from an individual superior device. For example, in an information system which adopts SCSI (Small Computer System Interface) as an interface between a superior device and a storage device, by use of a reserve series command which is defined by SCSI, it is possible to realize the shared exclusive control with a logical volume unit of the storage device. On the occasion that a certain superior device reserved the logical volume, realized is such a situation that read-access and write-access become possible only from the reserved superior device.

In the SCSI reserve series command, also prepared is an extension command by which the shared exclusive control can be carried out with a block unit of a disk. A SCSI command for reserving a partial region (extent) on this logical volume is defined as an extent reserve (hereinafter, referred to as "region reserve"). A region to be reserved has a reserve attribute. The reserve attribute enables read-share, exclusive-write, exclusive-read, and exclusive access. A technology regarding SCSI-2 is described in an item 6.15 of SCSI-2 DETAIL COMMENTARY, Volg.3 published by CQ Publishing Co., Ltd. on February 1, 1997 (Non-Patent Document 1).

Under a current remote copy technology environment, a shared exclusive control mechanism by the reserve series command

of SCSI is not considered, and even on the occasion that the logical volume in a certain storage device system was locked by the reserve command, the locked state is not transmitted to a remote copy correspondence logical volume in another storage device system.

There was such a problem that the remote copy of the above-described related art enables writing only into the copy-source logical volume from a superior device. Also, there was such a problem that the locked state by the reserve series command is not transmitted to the remote copy correspondence logical volume.

SUMMARY OF THE INVENTION

A first object of this invention is to control so as to realize bi-directional copying without fixing a copy direction to one direction between storage device systems which configure the copy pair.

A second object of this invention is, under the bi-directional copying, to propagate a situation reserved by the reserve series command between storage device systems in which the remote copy is carried out.

This invention, in order to enable the bi-directional copying between storage device systems, provides a data consistency holding control device in storage device systems which configure the copy pair. This data consistency holding

control device controls so as to write write-in data which was received from the superior device and write-in data which was received from the other storage device system through the communication interface device into the physical storage device after they are made to wait on a temporary storage device for more than predetermined time from the reception time corresponding to the write-in data to the logical volume which forms the copy pair, so that, when write-in data which was received from the superior device and write-in data which was received from the other storage device system through the communication interface device are written in an overlapped manner into the same storage location of the physical storage device, they are written in the order of the reception time when the write-in data was received from the superior device.

Also, a storage device system of this invention further has a device for receiving a request for locking a partial region of the logical volume from the superior device and for locking the partial region, a device for transmitting the locking request which was received through the communication interface device to the other storage device system, a device for receiving the locking request through the communication interface device from the other storage device system and for locking a partial region designated, and a device for rejecting a request of write-in data from the superior device and the other storage device system to the partial region, except such a case that it is a request

from the superior device in which the partial region was locked.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention, together with further advantages thereof, may best be understood by reference to the following description taken in conjunction with the accompanying drawings in which:

FIG.1 is a block diagram of a computer system of a mode for carrying out the invention;

FIG.2 is a view explaining a bi-directional remote copy of the mode for carrying out the invention;

FIG.3 is a hardware block diagram of a storage device system of the mode for carrying out the invention;

FIG.4 is a software block diagram of a storage device system of a first embodiment;

FIG.5 is a block diagram of a data consistency holding table of the first embodiment;

FIG.6 is a view showing an example of a temporary bit map table of the first embodiment;

FIG.7 is a flow chart showing process procedures for processing I/O which was received from a host I/F of the first embodiment;

FIG.8 is a flow chart showing process procedures for processing I/O which was received from DKCI/F of the first embodiment;

FIG.9 is a flow chart at the time of write-in request of

a data consistency holding control part of the first embodiment;

FIG.10 is a flow chart of a doubled block deletion process of the data consistency holding control part of the first embodiment;

FIG.11 is a flow chart of a simultaneous write-in process of the data consistency holding control part of the first embodiment;

FIG.12 is a software block diagram of a storage device system of a third embodiment;

FIG.13 is a view explaining bit map values of the third embodiment;

FIG.14 is a view showing an example of a locked state holding table of the third embodiment;

FIG.15 is a flow chart on the occasion of processing I/O which was received from a host I/F of the third embodiment;

FIG.16 is a flow chart showing process procedures for processing I/O which was received from DKCI/F of the third embodiment;

FIG.17 is a flow chart at the time of write-in request of a data consistency holding control part of the third embodiment;

FIG.18 is a flow chart at the time of write-in request of the data consistency holding control part of the third embodiment;

FIG.19 is a flow chart of a superior device usable/unusable

judging process of an exclusive control part of the third embodiment;

FIG.20 is a flow chart of a lock forming process by a lock forming request which was received from a host I/F of the exclusive control part of the third embodiment;

FIG.21 is a flow chart of a lock forming process by a lock forming request which was received from DKCI/F of the exclusive control part of the third embodiment;

FIG.22 is a block diagram of a computer system to which the third embodiment is applied;

FIG.23 is a process sequence chart regarding one example of processes of the system of FIG.22;

FIG.24 is a flow chart of a lock release process by a lock release request which was received from a host I/F of the exclusive control part of the third embodiment; and

FIG.25 is a flow chart of a lock release process by a lock release request which was received from DKCI/F of the exclusive control part of the third embodiment.

DESCRIPTION OF THE PREFERED EMBODIEMNTS

(1) First Embodiment

Hereinafter, a first embodiment which relates to bi-directional copying will be described with reference to the drawings.

FIG.1 is a block diagram of a computer system 1100 of this

mode for carrying out the invention. The computer system 1100 is configured by a site 1110 which comprises a plurality of superior devices 1000 and 1010 which were connected through SAN(Storage Area Network) 1040 to a storage device system 1070, and a site 1120 which comprises a plurality of superior devices 1020 and 1030 which were connected through SAN 1050 to a storage device system 1080. The storage device system 1070 and the storage device system 1080 are connected by a remote copy line 1060 which used a dedicated line or a public line. The storage device system 1070 and the storage device system 1080 can communicate to each other through the remote copy line 1060, using SCSI protocol.

FIG.2 is a view describing such a situation that bi-directional copying is being carried out in the computer system 1100 shown in FIG.1. In FIG.2, the superior device 1010 carries out write-in B 1200 of data to the storage device system 1070. The data of the write-in B 1200, after it was stored in the storage device system 1070, is transmitted (arrow 1240) through the remote copy line 1060, and copied by the storage device system 1080. In the same manner, as to write-in C 1210 from the superior device 1020 to the storage device system 1080, data is also transmitted (arrow 1230) through the remote copy line 1060 and copied by the storage device system 1070. In the same manner, as to write-in D 1220 from the superior device 1030 to the storage device system 1080, data is also transmitted (arrow

1250) through the remote copy line 1060 and copied by the storage device system 1070. In short, respective write-in B 1200, C 1210 and D 1220 of each superior device 1010, 1020 and 1030 are written into storage systems of respective sites, and then, copied by storage device systems of other part sites without intervention of the superior devices.

FIG.3 shows a hardware structure of the storage device system 1070. In passing, the storage device system 1070 may include a disk array device, a semiconductor storage device and so on. The storage device system 1070 is configured by a host I/F 1300, DKCI/F 1320, a disk control part 1350, a shared memory 1360, a cache memory 1340, a switching control part 1330 which is configured by a cross bus switch etc. for connecting them in such a manner that they can communicate with each other, a timer 1310, a physical disk 1370 and a processor 1380, and so on.

The host I/F 1300 has a CPU and a memory, and controls a communication with at least one superior device. The host I/F 1300 receives a data I/O request from the superior device and writes the data I/O request into the shared memory 1360. In passing, a function of remote copy is provided by executing a program by which the CPU of DKCI/F 1320 realizes this function.

The cache memory 1340 is mainly used for temporarily storing data which is given and received between the host I/F 1300, DKCI/F 1320 and the disk control part 1350. For example,

in case that a data input/output command which the host I/F 1300 received from a superior device is a write-in command, the host I/F 1300 writes the write-in data which was received from the superior device into the cache memory 1340. Also, the disk control part 1350 reads out the write-in data from the cache memory 1340 and writes it into the physical disk 1370.

The disk control part 1350 has a CPU and a memory, and reads out an I/O request which was written into the shared memory 1360 from the host I/F 1300 and DKCI/F 1320, and carries out processes such as write-in and read-out etc. of data to the physical disk 1370, in accordance with a command which was set in the I/O request (in this embodiment, a command of SCSI standard). The disk control part 1350, in case of a read-out command, writes data which was read out from the physical disk 1370, into the cache memory 1340. Also, a write-in completion notice and a read-out completion notice etc. of data are transmitted to the host I/F 1300. There is such a case that the disk control part 1350 has a function for dispersing and locating one logical volume in a plurality of physical disks 1370, in accordance with so-called RAID (Redundant Array of Inexpensive Disks) system RAID levels (e.g., 0, 1, 5).

The physical disk 1370 is a physical storage device such as for example, a hard disk device, for finally storing write-in data. The physical disk 1370 may be integrated with the storage device system and may be disposed in a separate housing. DKCI/F

1320 has a CPU and a memory, and is a communication interface for carrying out data transfer with another storage device system, and transfer of data to another storage device system in the remote copy is carried out through this DKCI/F 1320. It is supposed that each storage device system has one timer 1310, and each timer is adjusted to become the same time as much as possible. The timer 1310 is used for storing such time that the host I/F 1300 conducted reception of I/O from a superior device, and so on.

The processor 1380 has a CPU and a memory, and executes a program which will be described later of one other than the host I/F 1300, KC/F 1320 and the disc control part 1350.

FIG.4 is a view showing a software structure regarding a bi-directional remote copy of this mode for carrying out the invention. This structure has, in order to realize the bi-directional remote copy of the storage device system, each program of a main control 2020, a cache part 2050 and a data consistency holding control part 2040. These programs are executed by the processor 1380. Also, a bit map table 2030 is disposed on the memory.

The main control 2020 receives an input/output request from the host I/F 1300, and activates a data consistency holding control part 2040 and a cache part 2050, and returns a result of the input/output process to the host I/F 1300. Also, it controls handing-over of the input/output request through DKCI/F

1320.

The bit map table 2030 is a table which was bit-mapped so that one block on the physical disk which is an object of the remote copy corresponds to one bit. The bit map table 2030 has two states of a bit value 0 and a bit value 1. The value 0 represents such a situation that data of a block of a disk corresponding to the pertinent bit is not cached in the cache memory 1340. The value 1 represents such a situation that data of a block of a disk corresponding to the pertinent bit is cached in the cache memory 1340, and there is latest data on the cache.

The cache part 2050 is a program which carries out a process for writing the write-in data from the host I/F 1300 and DKCI/F 1320 into the cache memory 1340. As to the cache, there is a read-cache other than a cache for write-in data. A read-cache technology is such a technology that, on the occasion that a superior device reads data in a storage device system, in order to enable handing-over of read-data to the superior device quicker than reading reference data by directly accessing to the physical disk 1370, on the cache memory 1340, data which is accessed frequently is possessed as cache data. However, in this mode for carrying out the invention, in order to clarify features of this invention, the cache part 2050 is made to become one for carrying out only cache control of the write-in data.

The data consistency holding control part 2040 is activated when the input/output request is one which relates to the data

write-in, and controls in such a manner that consistency of data is held between storage device systems in which the bi-directional copying is carried out, on the basis of the bit map table 2030.

The write-in data from the host I/F 1300 and DKCI/F 1320 to the storage device system, after it was held for predetermined time on the cache memory 1340, is written through the disk control part 1350 into the physical disk 1370. The data consistency holding control part 2040 controls monitoring time for holding the write-in data in the cache. This time is made to become sufficiently longer time, considering transfer time during in which write-in data is transmitted to a copy destination and time for a control process in this invention.

In this embodiment, it is held on the cache memory 1340 for 3 minutes from such time that the host I/F 1300 received a write-in request from a superior device. The data consistency holding control part 2040 controls in such a manner that the write-in data which exists on the cache memory 1340 for more than 3 minutes is written into the physical disk 1370 simultaneously at almost the same time between storage device systems which configure the copy pair. This simultaneous write-in is carried out every 1 minute. For example, when time is 00 hour 00 minute 00 second, 00 hour 01 minute 00 second, and 00 hour 02 minute 00 second, the simultaneous write-in is carried out. In short, data which was cached on the cache memory

1340 for more than 3 minutes and less than 4 minutes is written into the physical disk 1370, on the occasion of a write-in process to a disk at time 00 second. Hereinafter, this write-in process to the physical disk 1370, which is carried out every 1 minute, is called as "simultaneous write-in process".

On the occasion of writing a write-in content into the physical disk 1370, the timers 1310 in the storage device systems of the copy pair are synchronized with each other so as to become the same time as much as possible, and it is controlled in such a manner that contents to be written into the physical disk 1370 become completely the same after lapse of 3 minutes. By this, each storage device system writes the same data into the physical disk 1370 at almost the same timing.

FIG.5 is a view showing a data format of a data consistency holding table 100 which is held by the data consistency holding control part 2040. The data consistency holding table 100 is configured by items of table control number 101, reception time 102, superior device identification number 103, object block start address 104, object size 105, storage serial number 106 and cache data storage address 107.

The reception time 102 stores such time that the host I/F 1300 received the write-in data from the superior device. The superior device identification number 103 is an identification number of a superior device by which write-in was carried out to a storage device. This identification number is unique with

respect to each superior device, and is IP address, WWN(World Wide Name) which is used as to a fiber channel, and so on. The object block start address 104 and the object size 105 are a block number of an object for write-in and the number of write-in blocks, respectively. The storage serial number 106 is a unique value which is given with respect to each storage device system, and becomes a value which represents which storage device system received a write-in request from a superior device. The cache data storage address 107 is an address of the cache memory 1340 in which the write-in data is stored. The cache data storage address 107, when a data storage region is secured on the cache memory 1340 by use of malloc etc. of the C language, indicates its head address (The deletion of data on the cache can be carried out by free etc. of the C language).

The data consistency holding table 100 is sorted in chronological order with the reception time 102 as the basis, and a latest record (or entry) is to come after a most tail end of a table. The table control number 101 is supposed to store integer control numbers from head of a table in such sequence of 1, 2, 3, ... The control number 1 is a record which represents the chronologically oldest table on the cache memory 1340, and records are arranged in such a manner that a record with the largest control number becomes a write-in record having the most recent time. Since a record which is next to the record with the largest control number is a unregistered entry, its control

number is substituted with -1.

Describing concretely by use of FIG.5, at present, 100 pieces of records are registered in the data consistency holding table 100. Each record is sorted in ascending sequence from the oldest one by the reception time 102. The table control number 101 is assigned in order up to 100 from 1 as a number of a head record. There exists no record which is one below 100 as the table control number 101. The table control number 101 of this unregistered record storage region is substituted with -1.

FIG.6 is a view showing a data format of the bit map table 2030. Each column of the bit map table 2030 is arranged in the order of a block, corresponding to a block on the physical disk. The bit value 1 or 0 is set in each column, and as above, it shows whether or not data of the block is cached. The temporary bit map table 200 is a table for storing a bit map with a scope of a block which is a target of an input/output of the input/output request from the superior device.

By use of FIGS.7 and 8, process procedures of the main control part 2020 will be described. The main control part 2020 processes the input/output request (hereinafter, referred to as I/O) which was received from the host I/F 1300 and DKCI/F 1320. I/O from the host I/F 1300 is I/O from the superior device, and I/O from DKCI/F 1320 is I/O from another storage device system. Here, a process regarding the request from the host I/F 1300

is shown in FIG.7, and a process regarding the request from DKCI/F 1320 is shown in FIG.8.

FIG.7 is a flow chart showing process procedures of the main control part 2020 which processes I/O from the host I/F 1300. When the I/O request has come from the host I/F 1300 to the storage device system, at a step 3000, discriminated is a reference series command (command such as read in SCSI) or an update change series command (command such as write in SCSI), and processes are bifurcated. In this embodiment, an attention is paid only to the reference series command and the update change series command. In case of the reference series command, a process goes to a step 3005. In case of the update change series command, a process goes to a step 3050.

In a step 3005, referred are bit values of the bit map table 2030 which correspond to a reference scope (reference block start address, the number of blocks) which is requested by the superior device, and they are temporarily copied in the bit map table 200. After the temporary bit map table 200 was prepared, a process goes to a step 3010. In the step 3010, it is judged whether the most recent content exists on the cache memory 1340 or exists on the physical disk 1370, in a block address which is indicated by the temporary bit map table 200.

In the reference scope which is requested by the superior device, a storage location of the most recent data is changed between a block scope of the bit value 0 and a block scope of

the bit vale 1 of the temporary bit map table 200. The reference request in the block scope of the bit value being 0 is in such a situation that data on the physical disk 1370 is the most recent data, and a process goes to a step 3020. The reference request in the block scope of the bit value being 1 is in such a situation that data on the cache memory 1340 is the most recent data, and a process goes to a step 3040. The step 3020 is a process for reading reference scope data through the disk control part 1350. The step 3040 reads the reference scope data through the cache part 2050. On the occasion of referring to data through the cache part 2050, by use of the data consistency holding table 100, data in a reference object scope is read. Concretely, as to records of the data consistency holding table 100, in the order from the largest one toward a smaller one of the table control number 101, the reference scope and the object block address 204, the object size 205 are compared, and an address of the cache memory 1340, in which changed data is cached is searched, and the reference scope data is read from the cache memory 1340.

After the step 3020 and the step 3040, in a step 3030, respective read-in data are joined as read-in data of the block scope of the temporary bit map table 200, and handed over to the host I/F 1300. After the data was handed over to the host I/F 1300, a process of I/O is completed. The host I/F 1300 transmits the read-in data which was handed over to the superior

device from which the reference request was issued.

The step 3050, steps 3060 and 3070 are processes on the occasion of having received the update change command from the host I/F 1300. On the occasion of carrying out write-in to data in the storage device system, the data consistency holding control part 2040 carries out processing, and therefore, the main control part 2020 carries out a process for handing over data to the data consistency holding control 2040, and for notifying completion of a process of data transmission to a remote copy destination storage device system and update change to the host. In case of data transmission to the remote copy destination storage device system, the main control part 2020 transmits, in addition to a content which was written, information of the reception time 102 which is an item of the data consistency holding table 200, the superior device identification number 103, the object block start address 104, the object size 105, and the storage serial number 206. Here, although it is not particularly described as to a transmission method, a vendor specific command which was formed by expanding a write command of SCSI may be used.

The step 3050 is a process of the data consistency holding control part 2040, and its detail will be described later. After the step 3050, process goes to the step 3060, and write-in data is handed over to DKCI/F 1320. DKCI/F 1320 transmits this write-in data to DKCI/F 1320 of a copy destination storage system.

After that, a process goes to the step 3070. The step 3070 carries out a write-in completion report to the host I/F 1300. The host I/F 1300 carries out notification of the write-in completion to the superior device which carried out the write-in request.

FIG.8 is a flow chart of a process when the main control part 2020 received the write-in data from DKCI/F 1320. A step 3100 is a data consistency holding control part process.

FIG.9 is a flow chart on the occasion that the data consistency holding control part 2040 processed the write-in request from the superior device. The data consistency holding control part 2040 is a control part which carries out a process of write data (write-in content), so that a logical volume for forming the remote copy pair becomes a volume which was completely geminated. The data consistency holding control part 2040 updates the data consistency holding table 100 and the bit map table 2030, and carries out a process for writing the write data on the cache memory 1340.

A step 3200 investigates whether or not there is a bit of the bit value 1 in a scope of the bit map table 2030 which corresponds to a write-in scope of the write data (block start address, and the number of blocks of a disk). As a result of this, on the occasion that the bit map in a scope of the write data is all the bit value 0, a process goes to a step 3270. In case that there is the bit value 1 even partially in the scope of the write data (in case that data of the same block as the

write-in block of the write data is cached on the cache), a process goes to a step 3210.

In the step 3210, the data consistency holding control part 2040 prepares the temporary bit map table 200 which has bits of the number of written blocks. The temporary bit map 200 corresponds to the write-in block of the write data, and a head of the temporary bit map 200 corresponds to a write-in head block of the write data. In the step 3210, referring to the bit map table 2030, a bit value in a scope of the write data is copied in the temporary bit map table 200 which was prepared. After the copy, a process goes to a step 3220.

The step 3220 searches a record which was written most recently of the data consistency holding table 100. In the search of the record which was written most recently, the data consistency holding table 100 is searched from a top in sequence, and an entry one above the entry in which the table control number became -1 is entered. The table control number 101 of record searched is put into a variable "current" on the memory. Current is made to become a variable which can be substituted with an integer. After the table control number was substituted in the current, a process goes to a step 3230.

Processes from the step 3230 to a step 3260 become a loop process, and a loop completion condition is that bit values of the temporary bit map table 200 become all 0. Hereinafter, procedures for making the bit values 0 will be described. The

write-in data of the temporary bit map table 200, in which the bit value corresponds to 1, to the block has already existed on the cache memory 1340. Thus, by detecting a record of the data consistency holding table 100, all of overlapped data (data of the bit value 1 in the temporary bit map table 200) on this cache memory 1340 are searched, and a bit value of the temporary bit map table 200, which corresponds to a scope of overlap of the record which was found in the search is made to become 0. Also, an overlapped block deletion process 3240 which will be described later is carried out in a scope of the overlapped data. By searching each record of this data consistency holding table 100, a process for making bit values of the temporary bit map table 200 to be 0 is carried out. When all of the bit values of the temporary bit map table 200 become 0, the loop process is completed.

The step 3230 judges whether or not there is a portion overlapped with a write-in scope of the write data, by use of the object block start address 104 and the object size 105 of a record in which the table control number 201 of the data consistency holding table 100 accords with variable current (hereinafter, referred to as "current record"). As a result of the judgment, in case that there is no overlapped portion, a process goes to a step 3250. In case that there is an overlap, a process goes to a step 3240. After the overlapped block deletion process of the step 3240 was completed, a process goes

to the step 3250. The step 3250 is a process for changing the current record which is a comparison object record to a record one above the same. Concretely, it is fine if a variable value of the current is made to become -1. After the value of the variable was changed, a process goes to a step 3260. The step 3260 is a process for carrying out a completion condition judgment of the loop process, and judges whether all of values of the temporary bit map table 200 become 0.

In case that all of the bit values of the temporary bit map table 200 become 0, a process goes to a step 3270. In case that 1 remains in the bit values, a process goes to the step 3230 which is a head of the loop. The step 3270 is a process for changing all of bit values of the bit map table 2030 which corresponds to the write-in scope of the write data to 1. After the process was completed, a process goes to a step 3280. In the step 3280, the data consistency holding control part 2040 writes information regarding the write data into the data consistency holding table 100. Concretely, a value of the table control number +1 of a most recently written record is written into the table control number 101. -1 is written into the table control number 101 one row below it, and furthermore, on the basis of information of the write data, the reception time 102, the superior device identification number 103, the object block start address 104, the object size 105, the storage serial number 106 and the cache data storage address 107 are written,

respectively. After the data consistency holding table 100 was updated, a process goes to a step 3290. The step 3290 writes the write data into an address which was set in the cache data storage address 107 on the cache memory 1340. After the write-in was completed, the process of the data consistency holding control part 2040 is completed.

FIG.10 is a flow chart showing an overlapped block deletion process. The overlapped block deletion process is a process which is carried out in case that there is an overlap between a write-in scope of the write data which is requested to be written into the storage device system and a write-in scope of the cache data which exists on the cache memory 1340. A fact that there is an overlap means that the write data overwrites the cache data or the cache data overwrites the write data. The overlapped block deletion process is a process for deleting a portion which is overwritten and updated in this overlapped scope.

(a) Case that the write data overwrites the cache data.

Normally, in such an environment that the remote copy is not used and so on, cache write-in is carried out in the order of the write data which was received by the host I/F 1300. Thus, the write data becomes the most recent data, and the write data overwrites the cache data.

(b) Case that the cache data overwrites the write data.

There is no such guarantee that the write data which was received by DKCI/F 1320 is the most recent write-in data. This

is write data which was transferred through the remote copy line 1060, after the write data was written into the remote copy destination storage device system, and time has already been passed since the reception time 102. Thus, there occurs such a case that the cache data becomes the most recent data rather than the write data. In this case, the cache data becomes the most recent data, and the cache data overwrites the write data.

In a step 3300, such times (reception times 102) that the write data and the cache data were received by the host I/F 1300 are compared. In case that the reception time of the write data is newer (more recent as time), a process goes to a step 3310. In case that the reception time of the cache data is newer, a process goes to a step 3350. On the occasion of comparing times in this embodiment, a process in case that completely the same times are compared is not included. Such a case that the same times are compared occurs in case that write-ins are applied to the same region simultaneously as a result of having measured in, for example, the storage device system 1070 and the storage device system 1120. As a process on this occasion, such an avoidance measure is conceivable that a priority storage device system has been determined in advance, and if the times are completely the same, write-in of the priority storage device system side is made to come first, and so on.

The step 3310 carries out a process for deleting an overlapped portion with a write-in area of the write data as

to the cache data which was compared. By this deletion process, record values of the deletion object data on the cache memory 1340 and the object block start address 104, the object size 105 and the cache data storage address 107 of the data consistency holding table 100 are changed.

Here, there is a case which requires an exceptional process. This exceptional process is carried out in case that, by a partial deletion of the cache data, an intermediate portion of the cache data is extracted, and the cache data is divided into two records. Concretely, it is supposed that the object size 105 is 20 blocks from the object block start address 104 of the cache data as a deletion object: 1000 address. In case that a write-in request of the object size 105:100 blocks came there from the write data object block start address 104:address 1020, a record which was written in the data consistency holding table 100 of the cache data, due to deletion of the overlapped portion with the write data, (the object block start address 104, the object size 105) is divided into two records of (1000,19), (1121, 79). A division process which is carried out here gets down a record after the record which is represented by current cache data by 1 row, and changes (+1) the control number to the record in which the table control number 101 was gotten down. By use of the record row which was prepared by this process and is 1 row below, the cache data is divided into two records.

A step 3320 judges whether all of the cache data are deleted

or not, as to the cache data which was changed in the step 3310. Such a situation that all are deleted is such a situation that a write-in scope of the cache data is completely included in a write-in scope of the write data, and all of the cache data are deleted. When a part of the cache data was deleted, a process goes to a step 3340. When all of the cache data were deleted, a process of a step 3330 is carried out. The step 3330 deletes a record row of the cache data of the data consistency holding table 100 which was compared, since all of the cache data which were compared were deleted. After the deletion, a record of 1 row behind and downward is shifted up by 1 row, and the table control number 101 of the shifted record is changed (-1).

A step 3350 carries out a process for deleting an overlapped portion with the write-in scope of the cache data which was compared with the write data. By this deletion process, size of the write data gets smaller, and it is to have new object block start address 104 and object size 105. In case that the write data was divided, a write-in process is applied to the divided write data as a separate write-in, respectively.

A step 3360 judges whether all of the write data are deleted or not, as to the write data which was changed in the step 3350. Such a situation that all of the write data were deleted is such a situation that a write-in scope of the write data is completely included in a write-in scope of the cache data, and all of the write data are deleted. When a part of the write data was deleted,

a process goes to a step 3340. When all of the write data were deleted, a process goes to a step 3370.

The step 3340 applies 0 to each of bit values of the temporary bit map table 200 which corresponds to the overlapped block processed in the step 3310 or the step 3350. After the change of the bit values, a process goes to the step 3250. The step 3370 completes a write-in process of the write data of the data consistency holding control part since all of the write data have been deleted and there is no content to be written.

FIG.11 is a flow chart showing a process (simultaneous write-in process) of the data consistency holding control part 2040 for writing the cache data on the cache memory 1340 to the physical disk 1370. The simultaneous write-in process is activated simultaneously in all storage device systems once every 60 seconds by the timer 1310. The step 3400 substitutes the variable current with 1 of the table control number 101 which is the oldest written record in the data consistency holding table 100. After the substitution, a process goes to the step 3410. In the step 3410, the table control number 101 refers to the reception time 102 of the record of the current. When the reception time 102 of the current record has not passed 3 minutes since the simultaneous write-in process start time, a process goes to the step 3470, and when more than 3 minutes pass, a process goes to a step 3420.

The step 3420 changes bit values of the bit map table 2030

which shows the write-in scope of the current record to all 0. After the change, a process goes to a step 3430. In the step 3430, the data consistency holding control part 2040 hands over the write-in data (data on the cache memory 1340 which was designated by the object block start address 104, the object size 105 and the cache data storage address 107) of the current record to the disk control part 1350. The disk control part 1350 writes the write-in content of the current record into the physical disk 1370. After data was handed over to the disk control part 1350, a process goes to a step 3440.

The step 3440 deletes the cache data on the cache memory 1340 which was handed over to the disk control part 1350. After the deletion, a process goes to a step 3450. The step 3450 carries out deletion of a record on the data consistency holding table 100 which is indicated by the current record. The deletion of the current record row is to delete information which is written into the table control number 101, the reception time 102, the superior device identification number 203, the object block start address 104, the object size 105, the storage serial number 106 and the cache data storage address 107 of the current record of the data consistency holding table 100, respectively. After the deletion of the deletion current row of the current record was completed, a process goes to a step 3460. The step 3460 applies +1 to the variable value of the current. The record which is indicated by the current becomes a record which was written

next to the record which was written in the physical disk 1370 this time. And, a process returns to the step 3410, and it is judged whether the current record is a simultaneous write-in process object record or not.

The step 3470 updates the data consistency holding table 100. The records close up upward in sequence so that a head record of the table does not become an unused record, and the table control number 101 is assigned again from an upper record like 1, 2, 3, ... in sequence. After the change of the data consistency holding table 100, the process is completed.

In passing, as a modified example of the first embodiment, such an embodiment that the bit map table 2030 is not provided is possible. In that case, the data consistency holding control part 2040 does not consider the overlapped block in the cache memory 1340, and stores the write-in data as it is in the order of the reception time. The data consistency holding control part 2040 carries out the write-in to the data consistency holding table 100 in the step 3280 and the write-in of the cache data in the step 3290. There is no overlapped block deletion process. Also, the steps 3010, 3020 and 3040 in case that the storage devices system received a reference series command from a superior device search the data consistency holding table 100 in sequence from the most recent record, and as to such a reference scope which the write-in data exists in the cache memory 1340, data is read out from the cache memory 1340, and as to such a

reference scope which the write-in data does not exist in the cache memory 1340, data is read out through the disk control part 1350 from the physical disk 1370.

(2) Second Embodiment

In the first embodiment, on the occasion of the write-in process of the write data, in the step 3060, the write-in content is handed over to DKCI/F 1320. Here, completion of the write-in process of the copy destination storage device system is not waited for. In the step 3070, consistency of data in the volumes of the copy pair changes, depending upon whether the write-in process completion of the copy destination storage device system is waited for, or is not waited for.

In the first embodiment, before the write-in process completion to the copy destination storage device system is completed, write-in completion to a superior device is notified. However, the write-in process to the copy destination storage device system requires data transfer time of the remote copy and time for carrying out an actual write-in process. During time which is required for this process, the write-in content differs between the copy destination storage device system and the copy source storage device system. In case that completely geminated volumes are formed as the remote copy pair, in the step 3070, the write-in completion notice of the copy destination storage device is waited for. In case of an environment with no influence even if consistency is lost with regard to data

which was written for time which is required for the write-in of the copy destination storage device system, the main control part 2020 notifies the write-in completion to the host I/F 1300, without waiting for the write-in completion of the copy destination storage device system.

(3) Third Embodiment

The third embodiment realizes propagation of reserve information by adding change to the second embodiment. The third embodiment is applied in case of carrying out an exclusive access of the disk region by the reserve of the region. FIG.12 shows a system structure of the third embodiment for propagating the reserve information of SCSI to the storage device system which is a target of the remote copy.

The system of FIG.12 has each program of an exclusive control part 4000, a data consistency holding control part 4020 and a main control 4030. These programs are executed by the processor 1380. Also, on a memory, a bit map table 4010 is provided. Other structural elements are the same as in the first embodiment.

The exclusive control part 4000 holds a lock state holding table 400, and manages a reserve state from a superior device. The exclusive control part 4000 controls in such a manner that each of the storage device systems which become objects of the bi-directional remote copy have the lock state holding tables 400 of the same content. By having this lock state holding table

400 of the same content, on the occasion that a superior device locked a volume of a certain storage device system, a pair volume of the other site which configures the copy pair is turned in a locked state.

The bit map table 4010 has 4 states which represent the bit map as compared to the bit map table 2030. By this, the bit map table 4010 holds the state of the disk by 2 bits, but not by the bit map of 1 bit to 1 block of the disk. Here, it is not "a bit map" which is normally used, but in this embodiment, it is called as a bit map.

In the data consistency holding control part 4020, there is a change of a bifurcation condition of a process due to the bit value by increase of the bit value which is taken by the bit map (increase of states which is possessed by the bit map) as compared to the data consistency holding control part 2040. In the main control part 4030, there is a change of a process due to such a fact that the host I/F 1300 and DKCI/F 1320 receive and process the reserve series command other than the reference command and the update change series command, as compared to the main control part 2020.

A bit map value table 300 of FIG.13 is a table for explaining a meaning of states which are represented by a bit value 0, bit value 1, bit value 2 and bit value 3 which are possessed by the bit map table 4010. The bit value 0 indicates such a state that it is not reserved in any superior device, and data of the physical

disk 1370 is the most recent one. The bit value 1 indicates such a state that it is not reserved in any superior device, and data of the cache memory 1340 is the most recent one. The bit value 2 indicates such a state that it is reserved in a certain superior device, and data of the physical disk 1370 is the most recent one. The bit value 3 indicates such a state that it is reserved in a certain superior device, and data of the cache memory 1340 is the most recent one.

FIG.14 shows a data format of a lock state holding table 400 which is held by an exclusive control part 4000. The lock state holding table 400 is configured by items of a control number 401, a lock start time 402, a superior device identification number 403, a lock object start address 404 and a lock object size 405. The lock start time 402 stores such time that a lock request was received from a superior device to the host I/F 1300 in the storage device system. The superior device identification number 403 is the same meaning as the superior device identification number 203 of the data consistency holding table 100. The lock start address 404 and the object size 405 set a block address and the number of blocks of a disk which becomes a lock object, respectively. The control number 401 is supposed to store integer control numbers from head of a table in such sequence of 1, 2, 3, ... The control number starts from 1 and increases with +1 at a time, and a control number next to a final record is set to -1.

FIGS. 15 and 16 are flow charts showing process procedures of the main control part 4030. FIG.15 is one in which a change was added to FIG.7, and FIG.16 is one in which a change was added to FIG.8. This changed portion will be hereinafter described. FIG.15 is a flow chart when the main control part 4030 received I/O from the host I/F 1300.

The main control part 4030, when it received an input/output request from the host I/F 1300, recognizes the reference series command, the update change series command and the lock series command (commands such as reserve, region reserve and release in SCSI) in a step 5000, and has a process bifurcated. On the occasion of the reference series command, a process goes to a step 5005. On the occasion of the update change series command, a process goes to a step 5010. On the occasion of the lock series command, a process goes to a step 5070. The steps 5010 and 5005 are the same process. This process investigates whether a process object scope in which reference and update change processes are carried out is locked or not by use of the reserve series command by another superior device. According to this lock state, the main control portion 4030 judges whether the superior device which issued reference and update change requests can process the reference and update change or not as to its requested scope. A detail of this judgment process will be described later.

In case of the reference series command, after the process

of the step 5005 was completed, a process goes to judgment result bifurcation of a step 5020. In case of the update change series command, after the process of the step 5010 was completed, a process goes to judgment result bifurcation of a step 5050. On the occasion of the reference series request, in case that a request scope can be referred in a step 5020, a process goes to a step 5030. In case that the reference is impossible, a process goes to a step 5040. On the occasion of the update change request, in case that a request scope can be updated and changed in the step 5050, a process goes to a step 5060. In case that the update change is impossible, a process goes to the step 5040. The step 5040 notifies, on the occasion that the reference/update change are impossible, unavailability of the request scope through the host I/F 1300 to a superior device (In the SCSI protocol, Reservation Conflict is returned to the superior device).

The step 5030 carries out processes of the steps 3005, 3010, 3020, 3040 and 3030. Here, in the bit map table 4010, block reference request of the bit values 0 and 2 is such a state that data on the physical disk 1370 is the most recent data, and reference data becomes one which is read in from the physical disk 1370. Reference requests of block scopes of the bit values 1 and 3 are such a state that data on the cache memory 1340 is the most recent data, and reference data becomes one which is read in from the cache memory 1340.

The step 5060 is a process of the data consistency holding

control part 4020 which will be described later. The step 5070 is a process of the exclusive control part 4000 on the occasion that a lock series command was handed over from the host I/F 1300 to the main control part 4030. The exclusive control part 4000 which will be described later gives back a return value to the main control part 4030. After reception of the return value was completed, a process goes to a step 5080. The step 5080 hands over the return value which was received from the exclusive control part 4000, to the host I/F 1300. After the return value was transmitted to the host I/F 1300, the process on the occasion of having received the lock series command is completed.

FIG.16 is a flow chart showing process procedures in case that the main control part 4030 received I/O from DKCI/F 1320. A step 5100 recognizes the update change series command or the lock series command, and has a process bifurcated. In case of the update change series command, a process goes to a step 5110. In case of the lock series command, a process goes to a step 5120. The step 5110 is a process of the data consistency holding control part 4020 which will be described later. The step 5120 is a step in case that the lock series command was handed over from DKCI/F 1320. A process of the step 5120 is carried out in the exclusive control part 4000, and the main control part 4030 receives its return value from the exclusive control part 4000. After reception of the return value was completed, a

process goes to a step 5130. The step 5130 hands over the return value which was received from the exclusive control part 4000, to DKCI/F 1320. After the return value was transmitted to DKCI/F 1320, the process is completed.

FIGS.17 and 18 are flow charts of the data consistency holding control part 4020. A changed point from the data consistency holding control part 2040 is on a change due to increase of states which are possessed by the bit map table 4010. FIG.17 is a change of the flow chart shown in FIG.9. In Fig.17, the step 3200 was changed to a step 5200, and the step 3210 was changed to a step 5210, and the step 3270 was changed to a step 5220. The step 5200 compares a write-in scope of the write data with corresponding bit values of the bit map table 4010. As a result of this comparison, when all of the bit values of the scope of the write data are the bit value 0 or 2 (when the most recent data exists in the physical disk 1370), a process goes to a step 5220. As to the scope of the write data, when corresponding bit values of the bit map table 4010 include the bit value 1 or 3 (when the most recent value exists on the cache memory 1340), a process goes to a step 5210.

In the step 5210, prepared is the temporary bit map table 200 with the same size as the scope of the write data. The temporary bit map table 200 corresponds to a write-in block of the write data. The temporary bit map table 200 substitutes the bit value 1 in a block which represents the bit value 1 or

3, in the bit map table 4010, and substitutes the bit value 0 in a block which represents the bit value 0 or 2. In short, with a block position of the disk in which change data exists on the cache memory 1340 as the bit value 1, it is stored in the temporary bit map table 200.

The step 5220 is a process for changing a bit value of the bit map table 4010. In case that the bit value of the bit map table 4010 before the process is 0 or 1, it is changed to 1, and in case that the bit value is 2 or 3, it is changed to 3.

FIG.18 is one in which the step 3420 of writing cache data on the cache memory 1340 into the physical disk 1370 (simultaneous write-in process) was changed to a step 5300. The step 5300 is a process for changing bit values of the bit map table 4010, but in case that a bit value of the bit map table 4010 which is an object to be changed is the bit value 1, it processes to change it to the bit value 0, and in case of the bit value 3, it processes to change it to the bit value 2. This is a process for changing the bit value of the bit map due to disappearance of the cache data on the cache memory 1340 by the simultaneous write-in process.

FIGS.19, 20 and 21 are flow charts of processes regarding the exclusive control part 4000. FIG.19 shows a process for judging whether the write-in process in the write-in scope is possible, on the occasion that a write-in request of a superior

device came, i.e., whether the write-in scope has been already locked by one other than the superior device which issued the write-in request.

A step 5400 judges whether the request scope is locked or not by the superior device, by referring to the bit map table 4010 which corresponds to a process request scope (scope for which reference, update change are requested). In case that all of the request scope are the bit value 0 or 1, the process request scope is not locked by any superior device, and the exclusive control part 400 sets availability to the return value. In case that the bit value 2 or 3 is included in the process request scope, a process goes to a step 5410. In the step 5410, it is judged whether a superior device which locks the process request scope is a superior device which is issuing the process request or not, by referring to the record of the lock state holding table 400. As to this judgment, in the bit map table 4010 which corresponds to the process request scope, all of records of the lock state holding table 400 of the bit value 2 or 3 are searched, and it is judged whether the superior device identification number 403 with the searched record is a superior device which is requesting a process this time or not. On the occasion of locking due to the superior device which is requesting a process, the write data at this time can be processed, and availability is set to the return value. On the occasion of a process request by one other than the superior device which

is requesting a process, unavailability is set to the return value.

FIG.20 is a flow chart of the exclusive control part 4000 on the occasion that the host I/F 1300 received a lock request of a disk from a superior device. A step 5500 judges whether the lock request scope (block scope of a disk which is requested to be locked) has been locked by another superior device, by referring to bit values of the bit map table 4010. When all of bit values which correspond to the lock request scope are 0 or 1 (such a state that any superior device does not lock), a process goes to a step 5530. In case that 2 or 3 is included in the bit values which correspond to the lock request scope, it means that the lock request scope has been already locked by a certain superior device. On this occasion, in a step 5505, it is judged whether or not reserve of a lock designation scope is a reserve state of a superior device which is requesting a lock this time. On the occasion of a lock request from the superior device which is locking, lock completion is set to the return value.

On the occasion of a lock state of another superior device, fundamentally, lock failure is set to the return value. However, under a remote copy environment, there may be such a case that a lock deletion process is being executed by the other site of the copy pair, and a lock request was received during a period that its process result arrives at the own site. Thus, by the

process of a step 5510,, the lock request is transmitted to the copy destination storage device system. This return value is judged in a step 5520, and on the occasion that lock failure was returned from a storage device system of the other party's site, lock failure is set to the return value. On the occasion of having received lock completion from the step 5520, a process goes again to the step 5500, and the lock process is carried out from the beginning.

The steps 5505 and 5520 firstly judge whether there is a state locked by another superior device with reference to the bit map table 4010 in the own host. Next, the return value of the lock request to other site is viewed. Since this process transmits lock failure to a superior device, after the bit map of the own site is judged and it was judged that a lock is impossible, it is faster than time for a superior device to issue a lock request again, which is effective. A process for transmitting a lock request to other site and for waiting for the lock request becomes a process necessary for corresponding to time-lag which relates to propagation in reserve information propagation under the remote copy environment.

Here, concretely, an environment which is required for the process of the step 5520 will be described by use of a computer system 6000 of FIGS.22 and 23. FIG.22 is a block diagram of this computer system. In this system, a host A 6010 and a host B 6020 of a site 1110 and a site 1120 establish a cluster

environment. It is supposed that this cluster is managing a logical volume as a resource of the cluster. In the site 1110, there is a logical volume A 6040, and in the site 1120, there is a logical volume B 6050. The respective logical volumes are ones in which the copy pair is formed by the bi-directional remote copy. The both hosts are to be ones which handle these plurality of logical volumes as the same logical volume. The hosts A and B are to communicate by use of an IP(Internet Protocol) network 6030.

Next, an example of state transition and processing of FIG.23 will be described. In this cluster, there may be such a case that exclusive control of a disk is carried out by a reserve command of SCSI. For example, it is a cluster server of Microsoft Corporation, and so on. Here, when the host A 6010 of the site 1110 manages the logical volume A 6040 as a resource in case of off-line, the host A 6010 reserves the logical volume A 6040. Next, a work of the host A 6010 is to be transferred (fail-over) to the host B 6020, and the logical volume A 6040 which has been used is to be transferred to the logical volume B 6050 which is the remote copy pair. The host A 6010 releases the logical volume A 6040, and the host B 6020 reserves the logical volume B 6050, and the host B 6020 carries out an operation by use of the logical volume B 6050. The hosts A 6010 and B 6020 communicate the fail-over by use of the IP network 6030, and carry out the transfer of the process, smoothly.

On this occasion, when the release of the host A 6010 and the reserve of the host B 6020 are carried out at almost the same time, in case that there is no processes of the steps 5510 and 5520, the host B 6020 can not reserve the logical volume B 6050. That is, regardless of such a fact that the logical volume A 6040 is released, the logical volume B 6050 is not released, and therefore, the host B 6020 can not reserve the logical volume B 6050. As above, due to such a fact that the released logical volume can not be reserved, there occurs a possibility of affecting to an operation of a cluster server. Thus, a process of not returning a result of the reserve request quickly to a reserve request host, but of waiting for a reply of the reserve request to the other party's site, like the steps 5510 and 5520, and of returning it to a host becomes necessary.

A step 5530 changes bit values of the bit map table 4010 which corresponds to the lock object region to 2 or 3. Concretely, in case that bit values of a change object bit map of the bit map table 4010 are 0, a process for changing to 2 is carried out, and in case that bit values are 1, a process for changing to 3 is carried out. After the bit values were changed, a process goes to a step 5540. The step 5540 registers records of the lock start time 402, the superior device identification number 403, the lock object start address 404 and the lock object size 405 of the lock request at such a record position that the control number 401 of the lock state holding table 400 is -1. The control

number 401 is set to such an integer that +1 is added to the control number 401 of a record one above. Next, the exclusive control part 4000 applies -1 to the control number 401 of a vacant record one below the registered record. After the write-in process of this record was completed, a process goes to a step 5550. The step 5550 transmits the lock request to the other party's site storage device system. Here, a transmission method is not particularly described, a vendor-specific command in which a write command of SCSI was expanded may be used.

In the processes of the steps 5510 and 5550, record items of the lock state holding table 400 and the return value at the time of the lock request are handed over through the remote copy line 1060 to DKCI/F 1320. The step 5550 judges the return value from the other party's site. On the occasion that lock completion is the return value, the lock completion is set to the return value. On the occasion that lock failure is the return value, the lock failure is set to the return value.

As a result of the processes of the steps 5550 and 5560, such a case that lock failure is returned from the other party's site will be described. This case is such time that the storage device system 1070 and the storage device system 1080 received the lock request from a superior device at almost the same time. In this case, when the lock request of the site 1110 failed, a lock of the site 1120 becomes successful. Then, the storage device system 1070 has to cancel the processes which were carried

out in the steps 5530 and 5540, but this cancel process is, as described later, carried out by a lock forming process of the storage device system 1080 which was successfully locked to the storage device system 1070 (see, the process of a step 5650 of FIG.21).

FIG.21 shows process procedures of the exclusive control part 4000 on the occasion that DKCI/F 1320 received the lock request from other site. A step 5600 judges bit values of a bit map which corresponds to the lock request scope in the bit map table 4010. In case that all of the lock request scope are the bit value 0 or 1, a process goes to the step 5530. In case that the bit value 2 or 3 is included in the lock request scope, a process goes to a step 5610. The step 5610 prepares the temporary bit map table 200 with a size corresponding to the lock request region, and, focusing on corresponding bit value 2, 3 of the bit map table 4010, bit map information is written into the temporary bit map table 200.

At the time of write-in of the bit map information, on the occasion of the bit value 2, 3 in the bit map table 4010, the bit value 1 is written into and on the occasion of the bit value 0, 1 in the bit map table 4010, the bit value 0 is written into the temporary bit map table 200. In this manner, the temporary bit map table 200 pays attention only to whether a corresponding block was locked or not, and a bit which corresponds to a block which was overlapped and locked becomes a bit value

having the bit value 1, and a bit which corresponds to a block which is not locked becomes a bit value having the bit value 0.

After the temporary bit map table 200 was prepared, a process goes to a step 5620. The step 5620 substitutes the control number 401 of a record which was written most recently (such a record one above that the control number 401 is -1) of the lock state holding table 400 in the variable current. After the substitution, a process goes to a step 5630.

The step 5630 judges whether or not there is a portion which is overlapped and locked in a scope of the bit value 1 of the temporary bit map table 200 and a lock scope of the current record. In case that it was overlapped, a bit value of the overlapped portion of the temporary bit map table 200 is made to become 0, and a process goes to a step 5640. In case of no overlap, a process goes to a step 5680. In the step 5680, the variable current is made to become -1, and a process goes to a step 5690. The step 5690 compares lock request time at this time with the lock start time 402 of the current record. In case that the lock start time 402 of the current record is older (time is earlier), since there is a block scope which was locked at older time than the lock request time at this time in a corresponding block scope, lock failure is set to the return value, and in case that the lock start time 402 of the current record is newer (time is later), a process goes to the step 5630.

A process in case of comparing completely the same time is similar to one in case of the step 3300.

In the step 5640, it is judged whether all of bit values of the temporary bit map table 200 become 0. On the occasion that all of the bit values become 0, it means that only a lock request, which is newer than that at the time of the lock request time of this time, comes in the corresponding block scope, and a process goes to a step 5650. On the occasion that 1 is included in the bit values, a process goes to a step 5680.

The step 5650, as to the record which was locked later than the lock request time and overlapped with the lock request scope, deletes the record from the lock state holding table 400. After the deletion, records of the lock state holding table 400 are arranged so as to eliminate a vacant record from a top of records of the lock state holding table 400, and control numbers are assigned again. As to bit values which represent the record scope to be deleted in the bit map table 4010, in case that they were the bit value 3, they are changed to the bit value 1, and in case that they were the bit value 2, they are changed to the bit value 0. Next, the steps 5530 and 5540 are carried out in sequence, and lock completion is set to the return value.

FIGS.24 and 25 shows processes on the occasion that the host I/F 1300 and DKCI/F 1320 received a lock release request (release command etc. of SCSI). FIG.24 shows a process on the occasion that the host I/F.1300 received the lock release request,

and FIG.25 shows a process on the occasion that DKCI/F 1320 received the lock release request from another site.

A step 5700, as to bit values of the bit map table 4010 which correspond to a lock release request scope, changes them to 0 when bit values before processing are 2, and changes them to 1 when bit values are 3. After the change, a process goes to a step 5710. The step 5710 deletes a record of the lock state holding table 400 which was lock-released. After the deletion, a record which was written one below the record which was lock-released is shifted up to one above, and after it was shifted, the control number 401 of the shifted record is assigned again. As to the control number 401, a control number of the shifted record may be made to become -1. A step 5720 is a process for transmitting the lock release request to DKCI/F 1320 of each storage device system of the copy pair. In each case of FIGS.24 and 25, after the lock release process was completed, lock release is set to the return value.

The lock release process of FIGS.24 and 25 does not separate various commands such as release, reset and so on of SCSI commands. In this embodiment, it is processed so that, on the occasion that the lock release came from all of superior devices, as to the lock release request scope, lock is released. In the processes of FIGS.24 and 25, if an identification judgment process etc. of a superior device is entered, it can be further adapted to a remote copy pair environment in which lock release

is carried out strictly in accordance with the SCSI protocol.

(4) Fourth Embodiment

The embodiment 3 realized propagation of reserve information on the occasion of having used the bi-directional remote copy of the embodiment 2. The embodiment 4 shows a method of realizing propagation of the reserve information on the occasion of having used the bi-directional remote copy of the embodiment 1.

In the bi-directional remote copy of the embodiment 1, a write-in content is handed over to DKCI/F 1320. Here, completion of the write-in process of the copy destination storage device system is not waited for. In this situation, there is a possibility that write-in is carried out during a period of the reserve process. With regard to the write-in content during a period of the reserve process, there is a necessity of selecting to handle a content which was written after reserve request time as the write-in content without any change, or to delete it as no existence of the write-in content. This process is to be entered before lock completion is set to the return value in FIGS.20 and 21.

Here, a process for deleting the write-in content and for establishing such a situation that the write-in was not carried out will be described. Before lock completion is set to the return value, it is searched whether or not there is write data which was written into the lock object scope of the data

consistency table after the lock start time, and when the write-in data exists, established such a situation that this write-in data was not written, and the record in the data consistency holding table 100 and the cache data on the cache memory 1340 were deleted. By this change, the propagation of reserve information on the occasion of having used the bi-directional remote copy of the embodiment 1 can be realized.

(5) Fifth Embodiment

The fifth embodiment changes the third embodiment, and realizes the propagation of reserve information by use of another means. In the fifth embodiment, the exclusive control part 400 in the storage device system does not have the lock state holding table 400, and manages the lock state by use of the protocol of SCSI, by transmitting the reserve series command directly to the physical disk 1370. At this time, reserve is carried out to the storage device system which is the other party of the copy pair through the remote copy line 1060, and the physical disk 1370 in the copy destination also carries out the reserve by use of SCSI ID of a superior device. In this reserve process, the disk control part 1350 locks the physical disk 1370 as reserve from a superior device which requested for locking, by use of third party reserve. On the occasion of requests such as reference, update change etc. from a superior device, firstly by use of the physical disk 1370, it is judged whether there is an available situation.

Since this embodiment used the SCSI protocol as it is, it is possible to precisely carry out the propagation of reserve information such as a reserve attribute etc. Here, in the reserve state, the disk is managed directly by a SCSI command, but data is managed by the data consistency holding control part 4020. Thus, there occurs such a situation that the physical disk 1370 has a precise reserve state but, does not have the most recent data.

According to this invention, when the copy pair is configured between a plurality of storage device systems, volumes which configure the copy pair can carry out copying in bi-directions. Each superior device can freely write in any volume which forms the copy pair. Also, it becomes possible to propagate the reserve state between the storage device systems under the bi-directional copying.

[FIG.1]

1000, 1010, 1020, 1030	SUPERIOR DEVICE
1060	REMOTE COPY LINE
1070, 1080	STORAGE DEVICE SYSTEM
1110, 1120	SITE
1100	COMPUTER SYSTEM

[FIG.2]

1000	SUPERIOR DEVICE (A)
1010	SUPERIOR DEVICE (B)
1020	SUPERIOR DEVICE (C)
1030	SUPERIOR DEVICE (D)
1200	WRITE-IN B
1210	WRITE-IN C
1220	WRITE-IN D
1060	REMOTE COPY LINE
1230	WRITE-IN C TO OTHER PARTY'S SITE SIDE
1240	WRITE-IN B TO OTHER PARTY'S SITE SIDE
1250	WRITE-IN D TO OTHER PARTY'S SITE SIDE
1070, 1080	STORAGE DEVICE SYSTEM

[FIG.3]

1000, 1010, 1020, 1030	SUPERIOR DEVICE
1060	REMOTE COPY LINE
1070, 1080	STORAGE DEVICE SYSTEM

1300	HOST I/F
1310	TIMER
1380	PROCESSOR
1330	SWITCHING CONTROL PART
1340	CACHE MEMORY
1350	DISK CONTROL PART
1360	SHARED MEMORY
1370	PHYSICAL DISK

[FIG.4]

1000, 1010, 1020, 1030	SUPERIOR DEVICE
1300	HOST I/F
2020	MAIN CONTROL PART
2030	BIT MAP
2040	DATA CONSISTENCY HOLDING CONTROL PART
2050	CACHE PART
1350	DISK CONTROL PART
1370	PHYSICAL DISK
1070, 1080	STORAGE DEVICE SYSTEM

[FIG.5]

100	DATA CONSISTENCY HOLDING TABLE
101	TABLE CONTROL NUMBER
102	RECEPTION TIME
103	SUPERIOR DEVICE IDENTIFICATION NUMBER

104	OBJECT BLOCK START ADDRESS
105	OBJECT SIZE
106	STORAGE SERIAL NUMBER
107	CACHE DATA STORAGE ADDRESS

[FIG.6]

2030	BIT MAP TABLE
200	TEMPORARY BIT MAP TABLE

[FIG.7]

ホストＩ／ＦからのＩ／Ｏ	I/O FROM HOST I/F
3000	REFERENCE SERIES COMMAND OR UPDATE CHANGE SERIES COMMAND
参照系のコマンド	REFERENCE SERIES COMMAND
3005	PREPARE TEMPORARY BIT MAP TABLE
3010	READ-OUT/BIFURCATION OF DATA
ビット値＝１の参照範囲	REFERENCE SCOPE OF BIT VALUE = 1
ビット値＝０の参照範囲	REFERENCE SCOPE OF BIT VALUE = 0
3020	DATA READ-OUT FROM DISK
3040	DATA READ-OUT FROM CACHE
3030	HAND OVER READ-OUT DATA TO HOST I/F
Ｉ／Ｏ処理終了	I/O PROCESS COMPLETION
更新変更系のコマンド	UPDATE CHANGE SERIES COMMAND
3050	DATA CONSISTENCY HOLDING CONTROL PART PROCESS

3060 HAND OVER WRITE-IN CONTENT TO DKC I/F

3070 TRANSMIT WRITE-IN COMPLETION TO HOST I/F

[FIG.8]

ＤＫＣ　Ｉ／ＦからのＩ／Ｏ　I/O FROM DKC I/F

3100 DATA CONSISTENCY HOLDING CONTROL PART PROCESS

Ｉ／Ｏ処理終了 I/O PROCESS COMPLETION

[FIG.9]

データ一貫性保持制御部処理 DATA CONSISTENCY HOLDING CONTROL PART PROCESS

3200 BIT VALUE JUDGMENT OF BIT MAP

ライトデータの範囲のビット値において１がある

THERE IS 1 IN BIT VALUES OF WRITE DATA SCOPE

3210 PREPARE TEMPORARY BIT MAP TABLE

3220 OBTAIN TABLE CONTROL NUMBER 101 OF RECORD WHICH WAS REGISTERED MOST RECENTLY

ビット値１がある THERE IS BIT VALUE 1

3230 OBJECT BLOCK SCOPE COMPARISON

対象ブロックの重なり有り THERE IS OVERLAP OF OBJECT BLOCKS

対象ブロックの重なり無し TEHRE IS NO OVERLAP OF OBJECT BLOCKS

3240 OVERLAPPED BLOCK DELETION PROCESS

3250 SUCH RECORD THAT CONTROL NUMBER OF COMPARISON OBJECT RECORD IS -1 IS MADE TO BECOME COMPARISON OBJECT RECORD

ライトデータ範囲のビット値がすべて０

ALL OF BIT VALUES OF WRITE DATA SCOPE ARE 0

3260 ALL OF BIT VALUES OF TEMPORARY BIT MAP 200 WERE 0?

すべて０ ALL 0

3270 BIT VALUES OF BIT MAP ARE MADE TO BECOME 1

3280 WRITTEN INTO DATA CONSISTENCY TABLE

3290 WRITE-IN OF CACHE DATA

Ｉ／Ｏ処理終了 I/O PROCESS COMPLETION

[FIG.10]

重複ブロック削除処理（ステップ３２３０）

OVERLAPPED BLOCK DELETION PROCESS (STEP 3230)

ライトデータの方が新しい WRITE DATA IS NEWER

キャッシュデータの方が新しい CACHE DATA IS NEWER

3300 RECEPTION TIMES OF WRITE DATA AND CACHE DATA ARE COMPARED

3310 PORTION WHICH WAS OVERLAPPED WITH WRITE-IN BLOCK OF WRITE DATA IS DELETED FROM CACHE DATA

3320 NEW CACHE DATA CONFIRMATION

キャッシュデータが一部削除された CACHE DATA IS PARTIALLY DELETED

キャッシュデータがすべて削除された CACHE DATA IS ALL DELETED

3330 DATA CONSISTENCY HOLDING TABLE 100 UPDATE PROCESS

3340 TEMPORARY BIT MAP UPDATE

ステップ３２５０へ TO STEP 3250

3350 PORTION WHICH WAS OVERLAPPED WITH WRITE-IN BLOCK OF CACHE DATA IS DELETED FROM WRITE DATA

3360 NEW WRITE DATA CONFIRMATION

ライトデータが一部削除された WRITE DATA IS PARTIALLY DELETED

ライトデータがすべて削除された WRITE DATA IS ALL DELETED

3370 PROCESS OF DATA CONSISTENCY HOLDING CONTROL PART IS COMPLETED

処理終了 PROCESS COMPLETION

[FIG.11]

データ一貫性保持制御部 DATA CONSISTENCY HOLDING CONTROL PART

物理ディスク書き込み処理 PHYSICAL DISK WRITE-IN PROCESS

3400 TABLE CONTROL NUMBER 1 IS MADE TO BECOME COMPARISON OBJECT RECORD

3410 RECEPTION TIME OF COMPARISON OBJECT RECORD IS TIME BEFORE MORE THAN 3 MINUTES?

3分以上前 BEFORE MORE THAN 3 MINUTES

3分経っていない 3 MINUTES HAVE NOT YET PASSED

3460 RECORD OF CONTROL NUMBER +1 OF COMPARISON OBJECT RECORD IS MADE TO BECOME COMPARISON OBJECT RECORD

3470 DATA CONSISTENCY HOLDING PART TABLE UPDATE

3420 CHANGE TO BIT VALUE 0 OF BIT MAP

処理終了 PROCESS COMPLETION

3430 CACHE DATA IS TRANSMITTED TO DISK CONTROL PART, AND WRITTEN INTO DISK

3440 CACHE DATA IS DELETED

3450 OBJECT RECORD DELETION OF DATA CONSISTENCY HOLDING TABLE

[FIG.12]

1000, 1010, 1020, 1030 SUPERIOR DEVICE

1300 HOST I/F

1060 REMOTE COPY LINE

4030 MAIN CONTROL PART

4010 BIT MAP

4000 EXCLUSIVE CONTROL PART

4020 DATA CONSISTENCY HOLDING CONTROL PART

2050 CACHE PART

1350 DISK CONTROL PART

1370 PHYSICAL DISK

1070, 1080 STORAGE DEVICE SYSTEM

[FIG.13]

300 BIT MAP VALUE TABLE

BIT VALUE	EXPLANATION
0	DISK IS NOT RESERVED DISK IS IN THE MOST RECENT STATE
1	DISK IS NOT RESERVED DATA WHICH IS NOT UPDATED IN DISK EXISTS ON CACHE
2	DISK IS RESERVED IN A CERTAIN SUPERIOR DEVICE, DISK IS IN THE MOST RECENT STATE
3	DISK IS RESERVED IN A CERTAIN SUPERIOR DEVICE, DATA WHICH IS NOT UPDATED IN DISK EXISTS ON CACHE

[FIG.14]

400 LOCK STATE HOLDING TABLE

401 CONTROL NUMBER

402 LOCK START TIME

403 SUPERIOR DEVICE IDENTIFICATION NUMBER

404 LOCK OBJECT START ADDRESS

405 LOCK OBJECT SIZE

[FIG.15]

ホストI／OからのI／O I/O FROM HOST I/O

5000 PROCESS/BIFURCATION WITH RESPECT TO EACH SCSI COMMAND

参照系のコマンド REFERENCE SERIES COMMAND

5005 EXCLUSIVE CONTROL PART LOCK JUDGMENT PROCESS (JUDGED WHETEHR IT IS AVAILABLE)

参照要求範囲が参照可能

REFERENCE REQEUST SCOPE CAN BE REFERRED

参照要求範囲が参照不可

REFERENCE REQUEST SCOPE CAN NOT BE REFERRED

5020 JUDGMENT RESULT BIFURCATION

5030 PROCESSES OF STEPS 3005, 3010, 3020, 3040 AND 3030

5040 TRANSMIT SUCH A FACT THAT OTHER SUPERIOR DEVICE HAS BEEN ALREADY USED TO HOST I/F

I／O処理終了 I/O PROCESS COMPLETION

更新変更系のコマンド UPDATE CHANGE SERIES COMMAND

5010 EXCLUSIVE CONTROL PART LOCK JUDGMENT PROCESS (JUDGED WHETEHR IT IS AVAILABLE)

更新変更要求範囲が更新変更不可

UPDATE CHANGE REQEUST SCOPE CAN NOT BE UPDATE-CHANGED

更新変更要求範囲が更新変更可能

UPDATE CHANGE REQUEST SCOPE CAN BE UPDATE-CHANGED

5050 JUDGMENT RESULT BIFURCATION

5060 DATA CONSISTENCY HOLDING CONTROL PART PROCESS

3060 WRITE-IN CONTENT IS HANDED OVER TO I/F BETWEEN DKC

3070 WRITE-IN COMPLETION IS TRANSMITTED TO HOST I/F

ロック系のコマンド LOCK SERIES COMMAND

5070 EXCLUSIVE CONTROL PART LOCK (FORM/RELEASE) PROCESS

5080 RETURN VALUE OF EXCLUSIVE CONTROL PART LOCK PROCESS IS TRANSMITTED TO HOST I/F

[FIG.16]

ＤＫＣ Ｉ／ＦからのＩ／Ｏ I/O FROM DKC I/F

5100 UPDATE CHANGE SERIES/LOCK SERIES PROCESS BIFURCATION

更新変更系のコマンド UPDATE CHANGE SERIES COMMAND

5110 DATA CONSISTENCY HOLDING OCNTROL PART PROCESS

I／O処理終了 I/O PROCESS COMPLETION

ロック系のコマンド LOCK SERIES COMMAND

5120 EXCLUSIVE CONTROL PART LOCK (FORM/RELEASE) PROCESS

5130 RETURN VALUE OF EXCLUSIVE CONTROL PART LOCK PROCESS IS TRANSMITTED TO DKC I/F

[FIG.17]

データ一貫性保持制御部処理 DATA CONSISTENCY HOLDING CONTROL PART PROCESS

5200 BIT VALUE JUDGEMENT OF BIT MAP

ライトデータの範囲のビット値において１又は３がある

THERE IS 1 OR 3 IN BIT VALUES OF WRITE DATA SCOPE

5210 TEMPORARY BIT MAP TABLE IS PREPARED

3220 RECORD WHICH WAS REGISTERED MOST RECENTLY IS MADE TO BECOME COMPARISON OBJECT RECORD

０以外の値がある TEHRE IS VALUE OTHER THAN 0

3230 OBJECT BLOCK SCOPE COMPARISON

対象ブロックの重なり有り

THERE IS OVERLAP OF OBJECT BLOCK

対象ブロックの重なり無し

THERE IS NO OVERLAP OF OBJECT BLOCK

3240 OVERLAPPED BLOCK DELETION PROCESS

3250 SUCH RECORD THAT CONTROL NUMBER OF COMPARISON OBJECT RECORD IS -1 IS MADE TO BECOME COMPARISON OBJECT RECORD

ライトデータ範囲のビット値がすべて０又は２

ALL OF BIT VALUES OF WRITE DATA SCOPE ARE 0 OR 2

3260 ALL OF BIT VALUES OF TEMPORARY BIT MAP 200 BECAME 0?

すべて０ ALL 0

5220 STATE OF BIT MAP IS MADE TO BECOME 1 OR 3

3280 WRITTEN INTO DATA CONSISTENCY TABLE

3290 WRITE-IN OF CACHE DATA

Ｉ／Ｏ処理終了 I/O PROCESS COMPLETION

[FIG.18]

データ一貫性保持制御部 DATA CONSISNTECY HOLDING CONTROL PART

物理ディスク書き込み処理 PHYSICAL DISK WRITE-IN PROCESS

3400 TABLE CONTROL NUMBER 1 IS MADE TO BECOME COMPARISON OBJECT RECORD

3410 RECEPTION TIME OF COMPARISON OBJECT RECORD IS TIME BEFORE MORE THAN 3 MINUTES

３分経っていない 3 MINUTES HAS NOT YET PASSED

３分以上前 BEFORE MORE THAN 3 MINUTES

3470 DATA CONSISTENCY HOLDING PART TABLE UPDATE

5300 CHANGE BIT VALUES OF BIT MAP TO 0 OR 2

3460 RECORD OF CONTROL NUMBER +1 OF COMPARISON OBJECT RECORD IS MADE TO BECOME COMPARISON OBJECT RECORD

3430 CACHE DATA IS TRANSMITTED TO DISK CONTROL PART, AND WRITTEN INTO DISK

3440 CACHE DATA IS DELETED

3450 OBJECT RECORD DELETION OF DATA CONSISTENCY HOLDING TABLE

処理終了 PROCESS COMPLETION

[FIG.19]

処理要求上位装置が利用可能か不可能かを判定

JUDGE WHETHER PROCESS REQUEST SUPERIOR DEVICE IS AVAILABEL OR UNAVAILABLE

5400 BIT VALUE JUDGEMENT OF BIT MAP

ビット値＝2、3 BIT VALUE = 2, 3

5410 JUDGE WHETHER A SUPERIOR DEVICE WHICH IS LOCKING IS A SUPERIOR DEVICE WHICH IS REQUESTING FOR PROCESSING

処理要求上位装置以外の上位装置によって既に対象ブロックがロックされている BY A SUPERIOR DEVICE OTHER THAN THE PROCESS REQUEST SUPERIOR DEVICE, OBJECT BLOCK HAS BEEN ALREADY LOCKED

利用不可を返す RETURN UNAVAILABILITY

ビット値＝0、1 BIT VALUE=0, 1

ロックしている上記装置による処理要求

PROCESS REQUEST BY A SUPERIOR DEVICE WHICH IS LOCKING

利用可能を返す RETURN AVAILABILITY

[FIG.20]

ホストＩ／Ｆからのロック形成処理開始

LOCK FORMING PROCESS START FROM HOST I/F

5500 BIT VALUE JUDGEMENT OF BIT MAP

ライトデータの範囲のビット値において１又は３がある

THERE IS 1 OR 3 IN BIT VALUES OF WRITE DATA SCOPE

ロック完了を受信 RECEIVED LOCK COMPLETION

5505 LOCK REQUEST SCOPE IS LOCKED BY OWN SUPERIOR DEVICE

はい YES

ロック完了を返す RETURN LOCK COMPLETION

いいえ NO

5510 TRANSMIT LOCK REQUEST TO OTHER PARTY'S SITE

5520 RETURN VALUE FROM OTHER PARTY'S SITE

ロック失敗を受信 RECEIVE LOCK FAILURE

ロック失敗を返す RETURN LOCK FAILURE

ロック指定範囲のビット値がすべて０又は１

ALL OF BIT VALUES OF LOCK DESIGNATED SCOPE ARE 0 OR 1

5530 BIT VALUES OF BIT MAP ARE MADE TO BECOME 2 OR 3

5540 WRITTEN INTO LOCK STATE HOLDING TABLE

5550 TRANSMIT LOCK REQUEST TO OTHER PARTY'S SITE

5560 RETURN VALUE FROM OTHER PARTY'S SITE

ロック完了を受信 RECEIVE LOCK COMPLETION

ロック完了を返す RETURN LOCK COMPELTION

[FIG.21]

ＤＫＣＩ／Ｆからのロック形成処理開始

LOCK FORMING PROCESS START FROM DKC I/F

5600 BIT VALUE JUDGMENT OF BIT MAP

ロック指定範囲のビット値に２又は３がある

THERE IS 2 OR 3 IN BIT VALUES OF LOCK DESIGNATED SCOPE

5610 TEMPORARY BIT MAP TALBE IS PREPARED

5620 RECORD WHICH WAS REGISTERED MOST RECENTLY IS MADE TO BECOME COMPARISON OBJECT RECORD

比較対象レコードのロック要求時刻の方が最近

LOCK REQUEST TIME OF COMPARISON OBJECT RECORD IS MORE RECENTLY

5690 LOCK REQUEST TIEM COMPARISON

5630 OBJECT BLOCK SCOPE COMPARISON

対象ブロックの重なり有り（ビット値０にする）

TEHRE IS OVERLAP OF OBJECT BLOCK (MADE TO BECOME BIT VALUE 0)

比較対象レコードのロック要求時刻の方が古い

LOCK REQUEST TIME OF COMPARISON OBJECT RECORD IS OLDER

5640 ALL PLACES IN WHICH BIT MAP BIT VALUE WAS 2 OR 3 WERE FOUND

対象ブロック重なり無し TEHRE IS NO OVERLAPPED OBJECT BLOCK

見つかっていない NOT FOUND

ロック失敗を返す RETURN LOCK FAILURE

5680 RECORD OF CONTROL NUMBER -1 OF COMPARISON OBJECT RECORD IS MADE TO BECOME COMPARISON OBJECT RECORD

ロック指定範囲のビット値がすべて０又は１

ALL OF BIT VALUES OF LOCK DESIGNATED SCOPE ARE 0 OR 1

5650 BY RECORD WHICH WAS LOCKED AFTER LOCK REQUEST, RECORD IN WHICH LOCK SCOPE IS OVERLAPPED IS DELETED

見つかった FOUND

5530 BIT VALUES OF BIT MAP ARE MADE TO BECOME 2 OR 3

5540 WRITTEN INTO LOCK STATE HOLDING TABLE

ロック完了を返す RETURN LOCK COMPLETION

[FIG.22]

6030 IP NETWORK

6000 COMPUTER SYSTEM

6010 HOST A

6020 HOST B

6040 LOGICAL VOLUME A

1060 REMOTE COPY LINE

6050 LOGICAL VOLUME B

1110, 1120 SITE

[FIG.23]

6010 HOST A

6040 LOGICAL VOLUME A

6050 LOGICAL VOLUME B

6020 HOST B

論理ボリューム運用中 UNDER OPERATION OF LOGICAL VOLUME

ホストAリザーブ HOST A RESERVE

論理ボリュームの運用停止　STOP OPERATION OF LOGICAL VOLUME

リリース　RELEASE

リザーブ解消処理　RESERVE RELEASE PROCESS

リザーブ要求　RESERVE REQUEST

リザーブ失敗　RESERVE FAILURE

論理ボリュームを運用開始　START OPERATION OF LOGICAL VOLUME

論理ボリュームの運用不可

OPERATION OF LOGICAL VOLUME IS UNAVAILABLE

[FIG.24]

ホストＩ／Ｆからの解消処理開始

START RELEASE PROCESS FROM HOST I/F

5700 BIT VALUES OF BIT MAP ARE MADE TO BECOME 0 OR 1

5710 DELETE RECORD OF LOCK STATE HOLDING TABLE

5720 TRANSMIT LOCK RELEASE REQUEST TO OTHER PARTY'S SITE

ロック解除を返す　RETURN LOCK RELEASE

[FIG.25]

ＤＫＣＩ／Ｆからの解消処理開始

START RELEASE PROCEDURE FROM DKC I/F

5700 BIT VALUES OF BIT MAP ARE MADE TO BECOME 0 OR 1

5710 DELETE RECORD OF LOCK STATE HOLDING TABLE

ロック解除を返す　RETURN LOCK RELEASE